

#24

RECEIVED

- 50 -

AUG 17 2000

## SEQUENCE LISTING

TECH CENTER 1600/2900



## (1) GENERAL INFORMATION:

(i) APPLICANT: ASHIKARI, Toshihiko  
 TANAKA, Yoshikazu  
 FUJIWARA, Hiroyuki  
 NAKAO, Masahiro  
 FUKUI, Yuko  
 SAKAKIBARA, Keiko  
 MIZUTANI, Masako  
 KUSUMI, Takaaki

(ii) TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY

(iii) NUMBER OF SEQUENCES: 31

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
 (B) STREET: 1737 King Street, Suite 500  
 (C) CITY: Alexandria  
 (D) STATE: Virginia  
 (E) COUNTRY: United States  
 (F) ZIP: 22314-2756

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
 (B) COMPUTER: IBM PC compatible  
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/894,356  
 (B) FILING DATE: 18-AUG-1997  
 (C) CLASSIFICATION:

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: JP 7-67159  
 (B) FILING DATE: 17-FEB-1995

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: JP 7-196915  
 (B) FILING DATE: 29-JUN-1995

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: JP 8-46534  
 (B) FILING DATE: 30-JAN-1996

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: WO PCT/JP96/00348  
 (B) FILING DATE: 16-FEB-1996

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Meuth, Donna M.  
(B) REGISTRATION NUMBER: 36,607  
(C) REFERENCE/DOCKET NUMBER: 001560-308

(ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: (703) 836-6620  
(B) TELEFAX: (703) 836-2021

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1703 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Gentiana triflora var. japonica  
(F) TISSUE TYPE: petal

(vii) IMMEDIATE SOURCE:  
(A) LIBRARY: cDNA library  
(B) CLONE: pGAT4

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 6..1412

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCATT ATG GAG CAA ATC CAA ATG GTG AAG GTT CTT GAA AAA TGC CAA	47
Met Glu Gln Ile Gln Met Val Lys Val Leu Glu Lys Cys Gln	
1 5 10	
GTT ACA CCA CCA TCT GAC ACA ACA GAT GTC GAG TTA TCG CTA CCG GTA	95
Val Thr Pro Pro Ser Asp Thr Thr Asp Val Glu Leu Ser Leu Pro Val	
15 20 25 30	
ACA TTC TTC GAT ATC CCC TGG TTG CAC TTG AAT AAG ATG CAG TCC CTT	143
Thr Phe Asp Ile Pro Trp Leu His Leu Asn Lys Met Gln Ser Leu	
35 40 45	
CTG TTT TAC GAC TTT CCG TAC CCA AGA ACA CAT TTC TTG GAC ACT GTT	191
Leu Phe Tyr Asp Phe Pro Tyr Pro Arg Thr His Phe Leu Asp Thr Val	
50 55 60	

ATC CCT AAT CTT AAG GCC TCT TTG TCT CTC ACT CTA AAA CAC TAC GTT Ile Pro Asn Leu Lys Ala Ser Leu Ser Leu Thr Leu Lys His Tyr Val 65 70 75	239
CCG CTT AGC GGA AAT TTG TTG ATG CCG ATC AAA TCG GGC GAA ATG CCG Pro Leu Ser Gly Asn Leu Leu Met Pro Ile Lys Ser Gly Glu Met Pro 80 85 90	287
AAG TTT CAG TAC TCC CGT GAT GAG GGC GAC TCG ATA ACT TTG ATC GTT Lys Phe Gln Tyr Ser Arg Asp Glu Gly Asp Ser Ile Thr Leu Ile Val 95 100 105 110	335
GCG GAG TCT GAC CAG GAT TTT GAC TAC CTT AAA GGT CAT CAA CTG GTA Ala Glu Ser Asp Gln Asp Phe Asp Tyr Leu Lys Gly His Gln Leu Val 115 120 125	383
GAT TCC AAT GAT TTG CAT GGC CTT TTT TAT GTT ATG CCA CGG GTT ATA Asp Ser Asn Asp Leu His Glu Leu Phe Tyr Val Met Pro Arg Val Ile 130 135 140	431
AGG ACC ATG CAA GAC TAT AAA GTG ATC CCG CTC GTA GCC GTG CAA GTA Arg Thr Met Gln Asp Tyr Lys Val Ile Pro Leu Val Ala Val Gln Val 145 150 155	479
ACC GTT TTT CCT AAC CGT GGC ATA GCC GTG GCT CTG ACG GCA CAT CAT Thr Val Phe Pro Asn Arg Gly Ile Ala Val Ala Leu Thr Ala His His 160 165 170	527
TCA ATT GCA GAT GCT AAA AGT TTT GTA ATG TTC ATC AAT GCT TGG GCC Ser Ile Ala Asp Ala Lys Ser Phe Val Met Phe Ile Asn Ala Trp Ala 175 180 185 190	575
TAT ATT AAC AAA TTT GGG AAA GAC GCG GAC TTG TTG TCC GCG AAT CTT Tyr Ile Asn Lys Phe Gly Lys Asp Ala Asp Leu Leu Ser Ala Asn Leu 195 200 205	623
CTT CCA TCT TTC GAT AGA TCG ATA ATC AAA GAT CTG TAT GGC CTA GAG Leu Pro Ser Phe Asp Arg Ser Ile Ile Lys Asp Leu Tyr Gly Leu Glu 210 215 220	671
GAA ACA TTT TGG AAC GAA ATG CAA GAT GTT CTT GAA ATG TTC TCT AGA Glu Thr Phe Trp Asn Glu Met Gln Asp Val Leu Glu Met Phe Ser Arg 225 230 235	719
TTT GGA AGC AAA CCC CCT CGA TTC AAC AAG GTA CGA GCT ACA TAT GTC Phe Gly Ser Lys Pro Pro Arg Phe Asn Lys Val Arg Ala Thr Tyr Val 240 245 250	767
CTC TCC CTT GCT GAA ATC CAG AAG CTA AAG AAC AAA GTA CTG AAT CTC Leu Ser Leu Ala Glu Ile Gln Lys Leu Lys Asn Lys Val Leu Asn Leu 255 260 265 270	815
AGA GGA TCC GAA CCG ACA ATA CGT GTA ACG ACG TTC ACA ATG ACG TGT Arg Gly Ser Glu Pro Thr Ile Arg Val Thr Thr Phe Thr Met Thr Cys 275 280 285	863

GGA TAC GTA TGG ACA TGC ATG GTC AAA TCA AAA GAT GAC GTC GTC GTA TCA Gly Tyr Val Trp Thr Cys Met Val Lys Ser Lys Asp Asp Val Val Ser 290 295 300	911
GAG GAA TCA TCG AAC GAC GAA AAT GAG CTC GAG TAC TTC AGT TTT ACA Glu Glu Ser Ser Asn Asp Glu Asn Glu Leu Glu Tyr Phe Ser Phe Thr 305 310 315	959
GCG GAT TGC CGA GGA CTT CTG ACG CCC CCG TGT CCG CCT AAC TAC TTT Ala Asp Cys Arg Gly Leu Leu Thr Pro Pro Cys Pro Pro Asn Tyr Phe 320 325 330	1007
GGC AAC TGT CTT GCG TCA TGC GTT GCA AAA GCA ACA CAT AAA GAG TTA Gly Asn Cys Leu Ala Ser Cys Val Ala Lys Ala Thr His Lys Glu Leu 335 340 345 350	1055
GTT GGG GAT AAA GGG CTT CTT GTT GCA GTT GCA GCT ATT GGA GAA GCC Val Gly Asp Lys Gly Leu Leu Val Ala Val Ala Ala Ile Gly Glu Ala 355 360 365	1103
ATT GAA AAG AGG TTG CAC AAC GAA AAA GGC GTT CTT GCA GAT GCA AAA Ile Glu Lys Arg Leu His Asn Glu Lys Gly Val Leu Ala Asp Ala Lys 370 375 380	1151
ACT TGG TTA TCG GAA TCT AAT GGA ATC CCT TCA AAA AGA TTT CTC GGG Thr Trp Leu Ser Glu Ser Asn Gly Ile Pro Ser Lys Arg Phe Leu Gly 385 390 395	1199
ATT ACC GGA TCG CCT AAG TTC GAT TCG TAT GGT GTA GAT TTT GGA TGG Ile Thr Gly Ser Pro Lys Phe Asp Ser Tyr Gly Val Asp Phe Gly Trp 400 405 410	1247
GGA AAG CCT GCA AAA TTT GAC ATT ACC TCT GTT GAT TAT GCA GAA TTG Gly Lys Pro Ala Lys Phe Asp Ile Thr Ser Val Asp Tyr Ala Glu Leu 415 420 425 430	1295
ATT TAT GTG ATT CAG TCC AGG GAT TTT GAA AAA GGT GTG GAG ATT GGA Ile Tyr Val Ile Gln Ser Arg Asp Phe Glu Lys Gly Val Glu Ile Gly 435 440 445	1343
GTA TCA TTG CCT AAG ATT CAT ATG GAT GCA TTT GCA AAA ATC TTT GAA Val Ser Leu Pro Lys Ile His Met Asp Ala Phe Ala Lys Ile Phe Glu 450 455 460	1391
GAA GGC TTT TGC TCT TTG TCA TAGTCTCTTT AATAGAACCA TATTTGCTGC Glu Gly Phe Cys Ser Leu Ser 465	1442
AATAAAAGTAC CAAGTCCTTT AGTAACACTA CACCAAACCC TACTTCGAG GCGGGAACAC	1502
CACAAACGAGG TTCAATCACT AGAAGGTTGT ACTTCATAAA TTCCAGAGGT CGAATATACA	1562
CCGTTGTCCT CTGAAAAGTT GAACCTCACA CCTGACATGG TGTTACGATA GGTATTGTAT	1622
AATGCCATTA TATACTTCCA TAAAGTATCC TATGCAATAG AGAACATGTT ATGTGTTAAA	1682

AAAAAAAAAAAAA AAAAAAAAAAAA A

1703

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1622 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Gentiana triflora* va. *japonica*  
(F) TISSUE TYPE: petal

(vi) IMMEDIATE SOURCE:

(A) LIBRARY: cDNA library  
(B) CLONE: pGAT106

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 35.1471

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAACCATTGA ATCCAATTAA TCTGATTTAT TAAG ATG GCA GGA AAT TCC GAG 52  
Met Ala Gly Asn Ser Glu  
1 5

GAT ATC AAA GTT CTT GAG AAA TGC CGT GTT GCG CCA CCA CCG GAC GCC 100  
 Asp Ile Lys Val Leu Glu Lys Cys Arg Val Ala Pro Pro Pro Asp Ala  
 10 15 20

GTC GCC GAG TTT ACA GTC CCA CTG TCG TTT TTC GAC ATG CGA TGG TTG 148  
 Val Ala Glu Phe Thr Val Pro Leu Ser Phe Phe Asp Met Arg Trp Leu  
                  25                 30                 35

ATC TCT GAT GCA GAA CAC CAT CTG CAT TTC TAC AGA TTC CGC CAT CCT 196  
 Ile Ser Asp Ala Glu His His Leu His Phe Tyr Arg Phe Arg His Pro  
 40 45 50

TGT CCC AAC TCT AAA TTT ATC ATT TCA TCC ATT AAA TCG TCC CTT TCC 244  
 Cys Pro Asn Ser Lys Phe Ile Ile Ser Ser Ile Lys Ser Ser Leu Ser  
 55 60 65 70

CTT GTT CTC AAA CAC TTT CTT CCG TTA GCC GGG AAT TTG ATT TGG CCG 292  
 Leu Val Leu Lys His Phe Leu Pro Leu Ala Gly Asn Leu Ile Trp Pro  
 75 80 85

GTA GAT TCC TCC GAT AGA ATG CCG GAG TTG CGT TAC AAG AAA GGG GAC		340
Val Asp Ser Ser Asp Arg Met Pro Glu Leu Arg Tyr Lys Lys Gly Asp		
90	95	100
TCC GTT TCT TTA ACA ATT GCA GAA TCG AGC ATG GAT TTT GAT TAT CTC		388
Ser Val Ser Leu Thr Ile Ala Glu Ser Ser Met Asp Phe Asp Tyr Leu		
105	110	115
GCC GGA GAT CAT CAG AGG GAT TCT TAT AAA TTC AAC GAT TTG ATT CCG		436
Ala Gly Asp His Gln Arg Asp Ser Tyr Lys Phe Asn Asp Leu Ile Pro		
120	125	130
CAG CTG CCA GAA CCG ATT GTA ACC TCC GGC GAC GAA GTA TTA CCA CTT		484
Gln Leu Pro Glu Pro Ile Val Thr Ser Gly Asp Glu Val Leu Pro Leu		
135	140	145
150		
TTT GCT TTA CAG GTG ACG GTG TTC TCC AAC ACC GGT ATA TGC ATT GGA		532
Phe Ala Leu Gln Val Thr Val Phe Ser Asn Thr Gly Ile Cys Ile Gly		
155	160	165
CGC AAT CTT CAT CAA GTT CTT GGT GAT GCC AGT TCT TTT CTG CAT TTT		580
Arg Asn Leu His Gln Val Leu Gly Asp Ala Ser Ser Phe Leu His Phe		
170	175	180
AAT AAA TTA TGG GTT TTG GTT GAC AAA TCC AAT GGA GAT TCA TTA AAG		628
Asn Lys Leu Trp Val Leu Val Asp Lys Ser Asn Gly Asp Ser Leu Lys		
185	190	195
TTC CTT CCA CTT TCT CTA CCT ATG TAC GAC AGA TCT GTG GTG CAA		676
Phe Leu Pro Leu Ser Ser Leu Pro Met Tyr Asp Arg Ser Val Val Gln		
200	205	210
GAT CCA TTT CAT ATT CGT CGA AAA ATC TAC AAT GAA AGA AAA CTG CTC		724
Asp Pro Phe His Ile Arg Arg Lys Ile Tyr Asn Glu Arg Lys Leu Leu		
215	220	225
230		
AAA TCT CAG GGC ACA CCT ACT GTT CTA AAT CCA GCA ATT TCT AAA GAT		772
Lys Ser Gln Gly Thr Pro Thr Val Leu Asn Pro Ala Ile Ser Lys Asp		
235	240	245
GAA GTT CGA GCC ACC TTC ATC CTA CAC CCT ATT GAT ATC ATG AAG CTC		820
Glu Val Arg Ala Thr Phe Ile Leu His Pro Ile Asp Ile Met Lys Leu		
250	255	260
AAG AAA TTC ATT TCG TCA AAA AAT CGC AAC TTA ACC GGT AGT AGT AAT		868
Lys Lys Phe Ile Ser Ser Lys Asn Arg Asn Leu Thr Gly Ser Ser Asn		
265	270	275
TAT AAT CTG TCA ACT TTC ACG GTG ACA TCT GCA CTG ATC TGG ACA TGC		916
Tyr Asn Leu Ser Thr Phe Thr Val Thr Ser Ala Leu Ile Trp Thr Cys		
280	285	290
TTG TCG AAA TCA TTA GAC ACC GTC GTA AGA GAG AAG GTG GAA GAG GAT		964
Leu Ser Lys Ser Leu Asp Thr Val Val Arg Glu Lys Val Glu Glu Asp		
295	300	305
310		

AAA CAT GCA GCA AAC TTA TGT GCT TTC ATC AAC TGC CGA CAA CGT TTT Lys His Ala Ala Asn Leu Cys Ala Phe Ile Asn Cys Arg Gln Arg Phe 315 320 325	1012
GCT CCG CCG ATA CCT CAA AAT TAC TTT GGA AAT TGC ATA GTG CCT TGT Ala Pro Pro Ile Pro Gln Asn Tyr Phe Gly Asn Cys Ile Val Pro Cys 330 335 340	1060
ATG GTG GGA TCG ACT CAT GAG CAA CTT GTA GGA AAT GAA GGG TTG TCG Met Val Gly Ser Thr His Glu Gln Leu Val Gly Asn Glu Gly Leu Ser 345 350 355	1108
GTA GCT GCA ACC GCC ATC GGA GAT GCT ATC CAT AAG AGG TTA CAT GAC Val Ala Ala Thr Ala Ile Gly Asp Ala Ile His Lys Arg Leu His Asp 360 365 370	1156
TAC GAA GGA ATT CTG AGA GGA GAT TGG ATA TCG CCG CCC CGA TCA ACA Tyr Glu Gly Ile Leu Arg Gly Asp Trp Ile Ser Pro Pro Arg Ser Thr 375 380 385 390	1204
TCT GCG GCA CCA AGG TCG ACG CTC ATT TAT GTC GTT GGA TCC GCA CAA Ser Ala Ala Pro Arg Ser Thr Leu Ile Tyr Val Val Gly Ser Ala Gln 395 400 405	1252
CGC AAT GTG CAT GAT TTT GAT GCA GAT TTT GGT TGG GGA AAG CTT GAA Arg Asn Val His Asp Phe Asp Ala Asp Phe Gly Trp Gly Lys Leu Glu 410 415 420	1300
AAG CAT GAA TCT GTT TCA ACT AAT CCT TCG GCA ACA CTA ATT TTG ATC Lys His Glu Ser Val Ser Thr Asn Pro Ser Ala Thr Leu Ile Leu Ile 425 430 435	1348
TCT CGG TCC AGA AGA TTT AAA GGA GCA CTT GAG CTT GGC ATT TCT TTG Ser Arg Ser Arg Arg Phe Lys Gly Ala Leu Glu Leu Gly Ile Ser Leu 440 445 450	1396
CCT AAG AAT AGG ATG GAC GCA TTT GCC ACC ATT TTT ACG AAT TTC ATC Pro Lys Asn Arg Met Asp Ala Phe Ala Thr Ile Phe Thr Asn Phe Ile 455 460 465 470	1444
AAT AGT CTC CAT GTG AGG AGC CCT TTG TAAGAAAAAA GTGGTATCAA Asn Ser Leu His Val Arg Ser Pro Leu 475	1491
TGTATAAAAA AGACAGACAA GTTATGATGC AACAAATGTT TTAGGAGATT ACAATCCAT	1551
GGGAAGATGT ATCAAACCTCA TCTCTCTATA TATATATATT CAATTGTTTT AAAAAAAA	1611
AAAAAAAAA A	1622

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1605 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Petunia hybrida*  
(F) TISSUE TYPE: petal

(vi) IMMEDIATE SOURCE:

- (A) LIBRARY: cDNA library
- (B) CLONE: pPAT48

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 67..1410

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAA GAG ATA GAA GTT GCA GAT CTT ACT GAT GAA GAA GGC ACC ACC AAA Glu Glu Ile Glu Val Ala Asp Leu Thr Asp Glu Glu Gly Thr Thr Lys 115 120 125	444
TTG CAG GAC TTG ATT CCT TGT AAT AAA ATC TTG AAT TTG GAA GGG CTT Leu Gln Asp Leu Ile Pro Cys Asn Lys Ile Leu Asn Leu Glu Gly Leu 130 135 140	492
CAT CGC CCT CTT CTA GCT GTG CAG CTC ACC AAG CTC AAG GAC GGG CTC His Arg Pro Leu Leu Ala Val Gln Leu Thr Lys Leu Lys Asp Gly Leu 145 150 155	540
ACC ATG GGA TTA GCA TTT AAC CAT GCT GTG CTG GAT GGT ACT TCG ACG Thr Met Gly Leu Ala Phe Asn His Ala Val Leu Asp Gly Thr Ser Thr 160 165 170	588
TGG CAC TTT ATG ACC TCG TGG TCC GAG CTT TGC TGT GGG TCC ACC TCA Trp His Phe Met Thr Ser Trp Ser Glu Leu Cys Cys Gly Ser Thr Ser 175 180 185 190	636
ATT TCT GTC CCA CCA TTC CTT GAA CGA ACC AAG GCT CGT AAC ACT CGA Ile Ser Val Pro Pro Phe Leu Glu Arg Thr Lys Ala Arg Asn Thr Arg 195 200 205	684
GTC AAG CTC AAC CTC TCT CAA CCA TCA GAT GCA CCC GAA CAT GCT AAG Val Lys Leu Asn Leu Ser Gln Pro Ser Asp Ala Pro Glu His Ala Lys 210 215 220	732
TCA GCA ACC AAC GGT GAT GTC CCG GCC AAC GTA GAC CCA CCT CTT CGC Ser Ala Thr Asn Gly Asp Val Pro Ala Asn Val Asp Pro Pro Leu Arg 225 230 235	780
GAA AGA GTA TTC AAG TTC TCC GAG TTA GCA ATT GAC AAA ATC AAG TCA Glu Arg Val Phe Lys Phe Ser Glu Leu Ala Ile Asp Lys Ile Lys Ser 240 245 250	828
ACA GTC AAT GCC AAC TCA GGA GAG ACG CCA TTC TCC ACA TTC CAA TCA Thr Val Asn Ala Asn Ser Gly Glu Thr Pro Phe Ser Thr Phe Gln Ser 255 260 265 270	876
CTC TCC GCA CAC GTG TGG CTA GCC GTC ACA CGT GCG CGC CAA CTC AAG Leu Ser Ala His Val Trp Leu Ala Val Thr Arg Ala Arg Gln Leu Lys 275 280 285	924
CCC GAG GAC TAC ACT GTG TAC ACT GTG TTT GCT GAT TGC AGG AAA AGG Pro Glu Asp Tyr Thr Val Tyr Thr Val Phe Ala Asp Cys Arg Lys Arg 290 295 300	972
GTT GAT CCT CCA ATG CCA GAA AGT TAC TTC GGC AAC CTA ATT CAG GCA Val Asp Pro Pro Met Pro Glu Ser Tyr Phe Gly Asn Leu Ile Gln Ala 305 310 315	1020
ATT TTC ACA GTG ACC GCG GCA GGT TTG TTA CTA GCA AGC CCG ATC GAG Ile Phe Thr Val Thr Ala Ala Gly Leu Leu Leu Ala Ser Pro Ile Glu 320 325 330	1068

TTC GCT GGT GGG ATG ATA CAA CAA GCG ATC GTG AAG CAT GAC GCT AAG Phe Ala Gly Gly Met Ile Gln Gln Ala Ile Val Lys His Asp Ala Lys 335 340 345 350	1116
GCC ATT GAT GAA AGA AAC AAG GAG TGG GAG AGC AAC CCG AAG ATC TTT Ala Ile Asp Glu Arg Asn Lys Glu Trp Glu Ser Asn Pro Lys Ile Phe 355 360 365	1164
CAG TAC AAA GAT GCT GGA GTG AAC TGT GTT GCT GTT GGA AGT TCG CCA Gln Tyr Lys Asp Ala Gly Val Asn Cys Val Ala Val Gly Ser Ser Pro 370 375 380	1212
AGG TTC AAG GTT TAC GAC GTG GAT TTT GGA TGG GGA AAG CCA GAG AGT Arg Phe Lys Val Tyr Asp Val Asp Phe Gly Trp Gly Lys Pro Glu Ser 385 390 395	1260
GTG AGG AGT GGT TCG AAC AAT AGG TTT GAT GGA ATG GTG TAT TTG TAC Val Arg Ser Gly Ser Asn Asn Arg Phe Asp Gly Met Val Tyr Leu Tyr 400 405 410	1308
CAA GGC AAA AAT GGA GGA AGA AGC ATT GAT GTG GAG ATT AGT TTG GAA Gln Gly Lys Asn Gly Gly Arg Ser Ile Asp Val Glu Ile Ser Leu Glu 415 420 425 430	1356
GCA AAT GCT ATG GAG AGG TTG GAG AAA GAT AAA GAG TTC CTC ATG GAA Ala Asn Ala Met Glu Arg Leu Glu Lys Asp Lys Glu Phe Leu Met Glu 435 440 445	1404
ACT GCT TAATTTGCTT AGCTTGGACT CAACTGGCTA CACTTTATTT ATGAGCTGCT Thr Ala	1460
ATGACTCAC A TGCATGTATG TTTATTTTT TTGGAGGGGT TCTTCCCTT TATTGTTTC	1520
TATGTTTTT CTTTCTTGTA CGTTATGAAG AGAAACCGAG TATAAAGGAA TAATGTTTC	1580
AGTTATTAAA AAAAAAAA AAAAA	1605

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1479 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Perilla ocimoides*

(F) TISSUE TYPE: leaf

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: cDNA library
- (B) CLONE: pSAT208

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..1340

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CC GTG ATC GAA ACG TGT AGA GTT GGG CCG CCG CCG GAC TCG GTG GCG	47
Val Ile Glu Thr Cys Arg Val Gly Pro Pro Pro Asp Ser Val Ala	
1 5 10 15	
GAG CAA TCG GTG CCG CTC ACA TTC TTC GAC ATG ACG TGG CTG CAT TTT	95
Glu Gln Ser Val Pro Leu Thr Phe Phe Asp Met Thr Trp Leu His Phe	
20 25 30	
CAT CCC ATG CTT CAG CTC CTC TTC TAC GAA TTC CCT TGT TCC AAG CAA	143
His Pro Met Leu Gln Leu Leu Phe Tyr Glu Phe Pro Cys Ser Lys Gln	
35 40 45	
CAT TTT TCA GAA TCC ATC GTT CCA AAA CTC AAA CAA TCT CTC TCT AAA	191
His Phe Ser Glu Ser Ile Val Pro Lys Leu Lys Gln Ser Leu Ser Lys	
50 55 60	
ACT CTC ATA CAC TTC TTC CCT CTC TCA TGC AAT TTA ATC TAC CCT TCA	239
Thr Leu Ile His Phe Phe Pro Leu Ser Cys Asn Leu Ile Tyr Pro Ser	
65 70 75	
TCC CCG GAG AAA ATG CCG GAG TTT CGG TAT CTA TCC GGG GAC TCG GTT	287
Ser Pro Glu Lys Met Pro Glu Phe Arg Tyr Leu Ser Gly Asp Ser Val	
80 85 90 95	
TCT TTC ACC ATC GCA GAA TCT AGC GAC GAC TTC GAT GAT CTC GTC GGA	335
Ser Phe Thr Ile Ala Glu Ser Ser Asp Asp Phe Asp Asp Leu Val Gly	
100 105 110	
AAT CGT CCA GAA TCT CCC GTT AGG CTC TAC AAC TTT GTC CCT AAA TTG	383
Asn Arg Pro Glu Ser Pro Val Arg Leu Tyr Asn Phe Val Pro Lys Leu	
115 120 125	
CCG CCC ATT GTC GAA GAA TCC GAT AGA AAA CTC TTC CAA GTT TTC GCC	431
Pro Pro Ile Val Glu Glu Ser Asp Arg Lys Leu Phe Gln Val Phe Ala	
130 135 140	
GTG CAG GTG ACT CTT TTC CCA GGC CGA GGC GTC GGT ATT GGA ATA GCA	479
Val Gln Val Thr Leu Phe Pro Gly Arg Gly Val Gly Ile Gly Ile Ala	
145 150 155	
ACG CAT CAC ACC GTT AGC GAC GCC CCG TCG TTT CTC GCG TTT ATA ACG	527
Thr His His Thr Val Ser Asp Ala Pro Ser Phe Leu Ala Phe Ile Thr	

160	165	170	175	
GCT TGG TCT TCA ATG AGC AAA CAC ATT GAA AAT GAA GAT GAA GAT GAA				575
Ala Trp Ser Ser Met Ser Lys His Ile Glu Asn Glu Asp Glu Asp Glu				
180		185	190	
GAA TTT AAA TCT TTG CCA GTT TTC GAT AGA TCC GTC ATA AAA TAT CCG				623
Glu Phe Lys Ser Leu Pro Val Phe Asp Arg Ser Val Ile Lys Tyr Pro				
195		200	205	
ACG AAA TTT GAC TCC ATT TAT TGG AGA AAC GCG CTA AAA TTT CCT TTG				671
Thr Lys Phe Asp Ser Ile Tyr Trp Arg Asn Ala Leu Lys Phe Pro Leu				
210		215	220	
CAA TCT CGT CAT CCC TCA TTA CCG ACG GAC CGC ATT CGA ACC ACG TTC				719
Gln Ser Arg His Pro Ser Leu Pro Thr Asp Arg Ile Arg Thr Thr Phe				
225		230	235	
GTT TTC ACC CAA TCC AAA ATT AAG AAA TTG AAG GGT TGG ATT CAG TCC				767
Val Phe Thr Gln Ser Lys Ile Lys Lys Leu Lys Gly Trp Ile Gln Ser				
240		245	250	255
AGA GTT CCA AGT TTA GTC CAT CTC TCA TCT TTT GTA GCG ATT GCA GCT				815
Arg Val Pro Ser Leu Val His Leu Ser Ser Phe Val Ala Ile Ala Ala				
260		265	270	
TAT ATG TGG GCT GGC ATA ACG AAA TCA TTC ACA GCA GAT GAA GAC CAA				863
Tyr Met Trp Ala Gly Ile Thr Lys Ser Phe Thr Ala Asp Glu Asp Gln				
275		280	285	
GAC AAC GAG GAT GCA TTT TTC TTG ATT CCG GTC GAT CTA AGG CCA CGA				911
Asp Asn Glu Asp Ala Phe Phe Leu Ile Pro Val Asp Leu Arg Pro Arg				
290		295	300	
TTA GAT CCG CCG GTT CCT GAA AAT TAC TTC GGG AAC TGC TTA TCG TAC				959
Leu Asp Pro Pro Val Pro Glu Asn Tyr Phe Gly Asn Cys Leu Ser Tyr				
305		310	315	
GCG CTG CCG AGA ATG CGG CGG CGA GAG CTG GTG GGA GAG AAA GGG GTG				1007
Ala Leu Pro Arg Met Arg Arg Glu Leu Val Gly Glu Lys Gly Val				
320		325	330	335
TTT CTG GCA GCT GAG GTA ATC GCG GCG GAG ATA AAA AAA AGG ATC AAC				1055
Phe Leu Ala Ala Glu Val Ile Ala Ala Glu Ile Lys Lys Arg Ile Asn				
340		345	350	
GAC AAG AGA ATA TTA GAA ACG GTG GAG AAA TGG TCG CCG GAG ATT CGT				1103
Asp Lys Arg Ile Leu Glu Thr Val Glu Lys Trp Ser Pro Glu Ile Arg				
355		360	365	
AAA GCG TTG CAG AAA TCA TAT TTT TCG GTG GCA GGA TCG AGC AAG CTA				1151
Lys Ala Leu Gln Lys Ser Tyr Phe Ser Val Ala Gly Ser Ser Lys Leu				
370		375	380	

GAT CTT TAC GGT GCA GAT TTT GGA TGG GGG AAG GCG AGA AAG CAA GAA	1199		
Asp Leu Tyr Gly Ala Asp Phe Gly Trp Gly Lys Ala Arg Lys Gln Glu			
385	390	395	
ATA TTG TCG ATT GAT GGG GAG AAA TAT GCA ATG ACG CTT TGT AAA GCC	1247		
Ile Leu Ser Ile Asp Gly Glu Lys Tyr Ala Met Thr Leu Cys Lys Ala			
400	405	410	415
AGG GAT TTC GAA GGA GGA TTG GAG GTT TGC TTG TCT TTG CCT AAG GAC	1295		
Arg Asp Phe Glu Gly Gly Leu Glu Val Cys Leu Ser Leu Pro Lys Asp			
420	425	430	
AAA ATG GAT GCT TTT GCT GCT TAT TTT TCA CTG GGA ATT AAT GGT	1340		
Lys Met Asp Ala Phe Ala Ala Tyr Phe Ser Leu Gly Ile Asn Gly			
435	440	445	
TAATAATGT ATGTAATTAA ACTAATATTA TTATGTAACA ATTAATTAAG TGTTGAGTAA	1400		
CGTGAAGAAT AATCCCTATT ATATATTAT GATTGGTTC AAATAAAGTG TAAAGCCTCT	1460		
TGAAAAAAA AAAAAAAA	1479		

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1508 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Senecio cruentus*
  - (F) TISSUE TYPE: petal
- (vii) IMMEDIATE SOURCE:
  - (A) LIBRARY: cDNA library
  - (B) CLONE: pCAT8
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 3..1364
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TG AAC ATT CTC GAA CAT GCC CGA ATA TCG GCC CCC TCG GGC ACC ATC	47		
Asn Ile Leu Glu His Ala Arg Ile Ser Ala Pro Ser Gly Thr Ile			
1	5	10	15

GGC CAT CGC TCG TTA TCT CTT ACT TTC TTC GAC ATT ACT TGG CTA CTC Gly His Arg Ser Leu Ser Leu Thr Phe Phe Asp Ile Thr Trp Leu Leu	20	25	30	95
TTC CCT CCG GTC CAC CAT CTT TTC TAT GAC TTT CCA CAT TCT AAA Phe Pro Pro Val His His Leu Phe Phe Tyr Asp Phe Pro His Ser Lys	35	40	45	143
TCC CAT TTC ATG GAC ACT ATT GTT CCC AGG CTA AAA CAA TCT TTA TCG Ser His Phe Met Asp Thr Ile Val Pro Arg Leu Lys Gln Ser Leu Ser	50	55	60	191
GTC ACT CTT CAA CAT TTT TTC CCG TTT GCT AGT AAT TTG ATT GTA TTT Val Thr Leu Gln His Phe Phe Pro Phe Ala Ser Asn Leu Ile Val Phe	65	70	75	239
CCT AAC ACT GAT GGT TCG GGT TTT AAT AAA AAA CCA GAA ATA AAA CAC Pro Asn Thr Asp Gly Ser Gly Phe Asn Lys Lys Pro Glu Ile Lys His	80	85	90	287
GTT GAA GGT GAT TCT GTG GTT ACT TTT GCA GAA TGT TGT CTT GAC Val Glu Gly Asp Ser Val Val Val Thr Phe Ala Glu Cys Cys Leu Asp	100	105	110	335
TTT AAT AAT TTG ACA GGA AAT CAT CCT CGA AAA TGT GAA AAC TTT TAT Phe Asn Asn Leu Thr Gly Asn His Pro Arg Lys Cys Glu Asn Phe Tyr	115	120	125	383
CCA CTT GTA CCT TCA TTG GGA AAT GCA ATC AAA TTA TGT GAT TGC GTC Pro Leu Val Pro Ser Leu Gly Asn Ala Ile Lys Leu Cys Asp Cys Val	130	135	140	431
ACG GTC CCA CTT TTT TCA CTT CAA GTG ACG TTT TTT CCG GGC TCG GGT Thr Val Pro Leu Phe Ser Leu Gln Val Thr Phe Phe Pro Gly Ser Gly	145	150	155	479
ATA TCA CTA GGA ATG ACG AAT CAT CAT AGC CTT GGT GAC GCT AGC ACG Ile Ser Leu Gly Met Thr Asn His His Ser Leu Gly Asp Ala Ser Thr	160	165	170	527
CGG TTC AAC TTT TTG AAA GGG TGG ACT TCG ATT ATT CAA TCT GGT GTA Arg Phe Asn Phe Leu Lys Gly Trp Thr Ser Ile Ile Gln Ser Gly Val	180	185	190	575
GAT CGG TCT TTT TTA ACG AAA GGA TCT CCA CCG GTT TTT GAT AGA TTG Asp Arg Ser Phe Leu Thr Lys Gly Ser Pro Pro Val Phe Asp Arg Leu	195	200	205	623
ATT AAC ATC CCA CAT TTA GAT GAA AAT AAG TTG AGA CAT ACA AGG CTC Ile Asn Ile Pro His Leu Asp Glu Asn Lys Leu Arg His Thr Arg Leu	210	215	220	671
GAA AGT TTT TAT AAA CCT TCG AGC CTT GTT GGT CCC ACT GAT AAA GTT Glu Ser Phe Tyr Lys Pro Ser Ser Leu Val Gly Pro Thr Asp Lys Val	225	230	235	719

CGG TCA ACG TTT GTG TTG ACC CGA ACT AAT ATC AAT CTA CTA AAG AAA			767
Arg Ser Thr Phe Val Leu Thr Arg Thr Asn Ile Asn Leu Leu Lys Lys			
240	245	250	255
AAG GTC TTA ACC CAA GTG CCA AAC TTG GAG TAC ATG TCA TCT TTT ACG			815
Lys Val Leu Thr Gln Val Pro Asn Leu Glu Tyr Met Ser Ser Phe Thr			
260	265	270	
GTA ACT TGT GGT TAT ATA TGG AGT TGC ATA GCG AAA TCA CTC GTA AAA			863
Val Thr Cys Gly Tyr Ile Trp Ser Cys Ile Ala Lys Ser Leu Val Lys			
275	280	285	
ATA GGA GAA AGA AAG GGC GAA GAC GAG TTA GAA CAG TTC ATA ATC ACC			911
Ile Gly Glu Arg Lys Gly Glu Asp Glu Leu Glu Gln Phe Ile Ile Thr			
290	295	300	
ATT GAT TGT CGA TCT CGT CTT GAT CCA CCA ATT CCC ACA GCC TAC TTT			959
Ile Asp Cys Arg Ser Arg Leu Asp Pro Pro Ile Pro Thr Ala Tyr Phe			
305	310	315	
GGT AAC TGT GGT GCA CCA TGT GTC CCG ACC TTA AAA AAT GTC GTT TTG			1007
Gly Asn Cys Gly Ala Pro Cys Val Pro Thr Leu Lys Asn Val Val Leu			
320	325	330	335
ACT ACG GAA AAT GGG TAT GCA CTT GGT GCT AAA GTA ATT GGA GAG TCT			1055
Thr Thr Glu Asn Gly Tyr Ala Leu Gly Ala Lys Val Ile Gly Glu Ser			
340	345	350	
ATA TGC AAA ATG ATA TAT AAT AAG GAC GGA ATC TTG AAA GAT GCC GCG			1103
Ile Cys Lys Met Ile Tyr Asn Lys Asp Gly Ile Leu Lys Asp Ala Ala			
355	360	365	
AGA TGG CAT GAA CCT TTC ATG ATC CCG GCT AGG AAG ATT GGT GTT GCT			1151
Arg Trp His Glu Pro Phe Met Ile Pro Ala Arg Lys Ile Gly Val Ala			
370	375	380	
GGT ACA CCT AAG CTC AAC TTG TAC GAC TTT GAT TTT GGG TGG GGG AAG			1199
Gly Thr Pro Lys Leu Asn Leu Tyr Asp Phe Asp Phe Gly Trp Gly Lys			
385	390	395	
CGC ATA AAG TAT GAG ACT GTT TCA ATA GAC TAT AAT ACG TCG ATT TCT			1247
Arg Ile Lys Tyr Glu Thr Val Ser Ile Asp Tyr Asn Thr Ser Ile Ser			
400	405	410	415
ATA AAT GCA AGC AAA ACA TCA GCA CAA GAT CTT GAA ATT GGA TTG AGT			1295
Ile Asn Ala Ser Lys Thr Ser Ala Gln Asp Leu Glu Ile Gly Leu Ser			
420	425	430	
CTA CCG AGT ATG CAA ATG GAG GCG TTT TCT AGC ATC TTT GAT GAA GGA			1343
Leu Pro Ser Met Gln Met Glu Ala Phe Ser Ser Ile Phe Asp Glu Gly			
435	440	445	
TTA GAG AGT CAA GTT TCA TTG TAGATCATCG TCCCCTTTT GTGTGCATCA			1394
Leu Glu Ser Gln Val Ser Leu			
450			

AGTTTCTGTC GTTTTATGA GTTGCCACTG TTCTATTCTT TAAGTATACC TTTCGACTAT	1454
GTTTGAAAGA TGCAACGATA TAAAATGAAA AAAAAAAA AAAAAAAA AAAA	1508

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1522 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lavandula angustifolia
  - (F) TISSUE TYPE: petal
- (vii) IMMEDIATE SOURCE:
  - (A) LIBRARY: cDNA library
  - (B) CLONE: pLAT21
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..1352

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

NTG ACC ACC CTC CTC GAA TCC TCC CGA GTG GCG CCG CCT CCA GGC ACG	48
Xaa Thr Thr Leu Leu Glu Ser Ser Arg Val Ala Pro Pro Pro Gly Thr	
1 5 10 15	
GTG GCT GAG CAG TCA CTC CCG CTC ACC TTC TTC GAC ATG ACG TGG CTG	96
Val Ala Glu Gln Ser Leu Pro Leu Thr Phe Phe Asp Met Thr Trp Leu	
20 25 30	
CAT TTC CAC CCC ATG CTT CAG CTT CTC TTC TAC GAA CTC CCC TGT TCC	144
His Phe His Pro Met Leu Gln Leu Leu Phe Tyr Glu Leu Pro Cys Ser	
35 40 45	
AAA CCC GCC TTC CTC GAA ACC GTC GTT CCG AAA CTC AAA CAA TCC TTA	192
Lys Pro Ala Phe Leu Glu Thr Val Val Pro Lys Leu Lys Gln Ser Leu	
50 55 60	
TCT CTA ACC CTC AAA CAC TTC TTC CCC CTT TCA TGC AAT CTA ATC TAC	240
Ser Leu Thr Leu Lys His Phe Phe Pro Leu Ser Cys Asn Leu Ile Tyr	
65 70 75 80	

CCT CTA TCG CCG GAG AAA ATG CCG GAG TTC CGG TAT CAG AAC GGT GAC Pro Leu Ser Pro Glu Lys Met Pro Glu Phe Arg Tyr Gln Asn Gly Asp 85 90 95	288
TCG GTT TCT TTC ACG ATT ATG GAG TCT GTC GGA GAT CAT CCG CAT TCC Ser Val Ser Phe Thr Ile Met Glu Ser Val Gly Asp His Pro His Ser 100 105 110	336
GCT CAT AAA TAC TAC TGC TTT GCC CCT AGC GAC GAT TAT GAA GAT CTC Ala His Lys Tyr Tyr Cys Phe Ala Pro Ser Asp Asp Tyr Glu Asp Leu 115 120 125	384
CAG CTG CCG CCG ATA GTC GAG GAA TCT GAT CGG AAA TTG TTT CAA GTT Gln Leu Pro Pro Ile Val Glu Glu Ser Asp Arg Lys Leu Phe Gln Val 130 135 140	432
TTA GCC GTG CAA GTG ACT CTG TTT CCC GGT CGC GGG GTG TGC ATC GGA Leu Ala Val Gln Val Thr Leu Phe Pro Gly Arg Gly Val Cys Ile Gly 145 150 155 160	480
ATA ACG ACG CAC CAC ACC GTT AGC GAT GCT CCA TCG TTT GTA GGG TTT Ile Thr Thr His His Thr Val Ser Asp Ala Pro Ser Phe Val Gly Phe 165 170 175	528
ATG AAG AGT TGG GCT TCC ATC ACT AAA TTC GGA GGA GAT GAT GAA TTC Met Lys Ser Trp Ala Ser Ile Thr Lys Phe Gly Gly Asp Asp Glu Phe 180 185 190	576
TTG GAC GGA AAA GGT GAA TGT TTG CCG GTT TTC GAC CGA TCG CTC GTG Leu Asp Gly Lys Gly Glu Cys Leu Pro Val Phe Asp Arg Ser Leu Val 195 200 205	624
AAT TAT CCG CCT AAA TTG GAC ACA TAT TTA TGG AAC AAC GCG CAG AAA Asn Tyr Pro Pro Lys Leu Asp Thr Tyr Leu Trp Asn Asn Ala Gln Lys 210 215 220	672
CGT CCG TTG GAA TCG CAG CAT CCA TCT TTA CCG ACG GAT CGG ATT CGA Arg Pro Leu Glu Ser Gln His Pro Ser Leu Pro Thr Asp Arg Ile Arg 225 230 235 240	720
GCT ACC TAC CTT TTC ACC CAA TCT GAA ATT AAG AAA TTG AAG GGT TTG Ala Thr Tyr Leu Phe Thr Gln Ser Glu Ile Lys Lys Leu Lys Gly Leu 245 250 255	768
ATT CAG AGA AAA GCC CCA AAT GTA GTT AAT CTC TCT TCC TTC GTC GCG Ile Gln Arg Lys Ala Pro Asn Val Val Asn Leu Ser Ser Phe Val Ala 260 265 270	816
ATC GCA GCT TAT ATC TGG ACC GGC ATC GCC AAA TCG GTC GGA GAT TAC Ile Ala Ala Tyr Ile Trp Thr Gly Ile Ala Lys Ser Val Gly Asp Tyr 275 280 285	864
AAA GAC GTG GAT GAC GAC AAA CGC GCT TTC TTT TTA ATT CCG ATC GAT Lys Asp Val Asp Asp Asp Lys Arg Ala Phe Phe Leu Ile Pro Ile Asp 290 295 300	912

TTA AGG CCG CGT TTG GAT CCG CCG GCT CCG GGG AAC TAC TTC GGA AAC Leu Arg Pro Arg Leu Asp Pro Pro Ala Pro Gly Asn Tyr Phe Gly Asn 305 310 315 320	960
TGT CTA TCG TTT GCG ATG GCG AAG ATC CTG CGG CGG GAT TTG GTC GGA Cys Leu Ser Phe Ala Met Ala Lys Ile Leu Arg Arg Asp Leu Val Gly 325 330 335	1008
GAT GAA GGG GTG TTT CGG GCA GCT GAG GCG ATC GCG GCG GAA ATA GAG Asp Glu Gly Val Phe Arg Ala Ala Glu Ala Ile Ala Ala Glu Ile Glu 340 345 350	1056
AAG AGG ACG AGC GAC AAG AAG ATT CTA GAA ACT GTG GAG AAC TGG CCG Lys Arg Thr Ser Asp Lys Lys Ile Leu Glu Thr Val Glu Asn Trp Pro 355 360 365	1104
TCT GAG ATT CGC GAA GCC TTG CAA AAC TGT TAT TTC TCG GTG GCG GGA Ser Glu Ile Arg Glu Ala Leu Gln Asn Cys Tyr Phe Ser Val Ala Gly 370 375 380	1152
TCG AGC AGG CTT GAT CTT TAC GGC GCG GAT TTT GGA TGG GGT AAG GCG Ser Ser Arg Leu Asp Leu Tyr Gly Ala Asp Phe Gly Trp Gly Lys Ala 385 390 395 400	1200
GTG AAG CAA GAG ATA CTG TCG ATT GAT GGA GAG AAG TTT ACG ATG TCG Val Lys Gln Glu Ile Leu Ser Ile Asp Gly Glu Lys Phe Thr Met Ser 405 410 415	1248
TTG TGT AAA CCG AGG GAT GCT GCC GGA GGA TTG GAG GTT GGA TTG TCT Leu Cys Lys Pro Arg Asp Ala Ala Gly Gly Leu Glu Val Gly Leu Ser 420 425 430	1296
TTG CCA AAG GAG GAA TTG CAA GCT TTT GAT GAT TAT TTT GCG GAG GGA Leu Pro Lys Glu Glu Leu Gln Ala Phe Asp Asp Tyr Phe Ala Glu Gly 435 440 445	1344
ATA AAG GGT TGATTAATCA TTTAATCATG TATTATGAAG TTGGATGAAA Ile Lys Gly 450	1393
TCCTCTGTTT CATCTCTATT GTTTAACAA TAATTTTTT CCATTGAAC TTTTGAGTC	1453
AATAAAAAAA AAAAAAAA AAAAAAATG AAAAAACTCA GTTATTTTTT TTTTTTTTTT	1513
TTTTTTTTT	1522

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Arg Phe Leu Gly Ile Thr Gly Ser Pro Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ile His Met Asp Ala Phe Ala Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly Val Glu Ile Gly Val Ser Leu Pro Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ala Ser Leu Ser Leu Thr Leu Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

His Tyr Val Pro Leu Ser Gly Asn Leu Leu Met Pro Ile Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Val Arg Ala Thr Tyr Val Leu Ser Leu Ala Glu Ile Gln Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ile His Met Asp Ala Phe Ala Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Lys Ile His Met Asp Ala Phe Ala Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Lys Ile His Met Asp Ala Phe Ala  
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AARATHCAYA TGGAYGCNTT YGC

23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTCGAGTTTT TTTTTTTTTT TTT

23

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTCACCATGG AGCAAATCCA AATGGT

26

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGAGTCGCC TCATCAC

17

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AACAGCTATG ACCATG

16

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Asp Phe Gly Trp Gly Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GAYTTYGGNT GGGGNAA

17

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TGGCAACTGT CTTGCGTCAT G

21

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCATGTCAGG TGTGAGGTTTC AAC

23

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATCGTTTCGC ATGATTGAAC

20

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCAGAAGAAC TCGTCAAGAA

20

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 12..53

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGGATCCAAC A ATG GAG CAA ATC CAA ATG GTG GCC GTG ATC GAA ACG TGT  
Met Glu Gln Ile Gln Met Val Ala Val Ile Glu Thr Cys  
1 5 10

50

AGA

53

Arg

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTAAAACGAC GGCCAT

16

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 12..45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGGATCCAAC A ATG GAG CAA ATC CAA ATG GTG AAC ATT CTC GAA C  
Met Glu Gln Ile Gln Met Val Asn Ile Leu Glu  
15 20 25

45

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTCGGAGGAA TTCGGCACGA C

21

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 18..35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AGTCGGATCC AACAAATG ACC ACC CTC CTC GAA TCC  
Thr Thr Leu Leu Glu Ser  
15

35